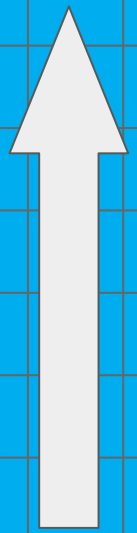
ML Academy

Register at
http://www.MLacademy.io



Overview of Machine Learning

Hands-on Learning Experience, Prepares you for a career in machine learning for your dream job.

Introductions

- Introduce yourself
  - Name
  - Education
- Share your expectation from this course
- If I am not afraid of failing I will……

A little about Venkat

- Masters in Com Sci & MBA
- Co-founded two different startups and successfully exited.
- Funded couple of startups that have raised Series-A funding
- Founded BI Engines (a BI Company)
- Currently in a early stages of IOT: ML product
- For the past two and a half years been teaching Machine Learning and Data Science in the Bay Area. Have taught over 1500 people!

What Can You Expect?

The workshop is meant to provide you with a base to build your machine learning skills. In particular you will learn to:

- Recognize problems that can be solved with Machine Learning
- Select the right technique (is it a classification problem? a regression? needs preprocessing?)
- Load and manipulate data with Panda
- Visualize and explore data with Seaborn
- Build regression models with Scikit-Learn
- Evaluate model performance with Scikit-Learn
- Solve one kaggle project.

What is Machine Learning?

- **Machine learning** is the art / science of programming computer so that they can learn from data
- Tom M. Mitchell provided a widely quoted,: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E."
- Due to the availability of large amounts of data (Big data), Machine learning has gained much importance in making data driven decisions, rather than hard coded responses.

Where is Machine Learning Used?

- Fraud detection.
- Web search results.
- Real-time ads on web pages
- Credit scoring and next-best offers.
- Prediction of equipment failures.
- New pricing models.
- Network intrusion detection.
- Recommendation Engines
- Customer Segmentation
- Text Sentiment Analysis
- Predicting Customer Churn
- Pattern and image recognition.
- Email spam filtering.
- Financial Modeling

Examples of Successful Machine Learning

- Spam filters….
- The heavily hyped, self-driving Google car? The essence of machine learning.
- Online recommendation offers such as those from Amazon and Netflix? Machine learning applications for everyday life.
- Knowing what customers are saying about you on Twitter? Machine learning combined with linguistic rule creation.
- Fraud detection? One of the more obvious, important uses in our world today.

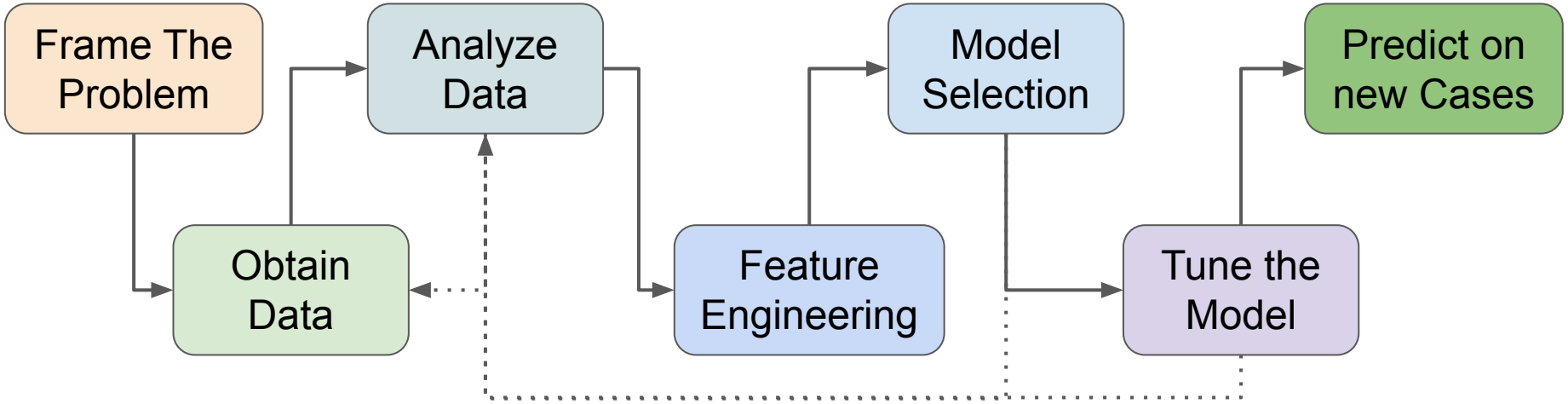
What is Needed to Learn ML?

- Computer science fundamentals
  - Data structures (stacks, queues, trees, graphs, etc.)
  - Algorithms (searching, sorting, optimization etc.)
  - Computability and complexity (Big-O notation).
- Probability and Statistics
  - Probability (conditional probability, Bayes rule, likelihood, independence, etc.).
  - Statistics (uniform, normal, binomial, poisson, etc.)
  - Analysis methods (Hypothesis testing, ANOVA, etc.)
  - College level Calculus and Linear algebra
  - Cheat sheets: Calculus, Linear Algebra and Statistics
- General Background
  - An inquisitive mind
  - Desire to learn something new

Environment Setup

- Create and Login to Gmaill
- Connect to Google Drive
- Connect Colab

End-to-End Supervised Machine Learning



ML Academy

SOLIVAR LABS

Machine Learning : What is Great For?

- Where existing problems require a lot hand tuning or lot of rules
  - ML can simplify code and perform better
- Complex problem for which there is no good solution
  - ML Techniques can find a solution
- Fluctuating environment
  - ML can adapt to change in data
- Getting insights about complex problems
  - ML can scan huge data problems

Types of Machine Learning Systems

- Whether or not they are trained with human supervision
  - Supervised, UnSupervised, SemiSupervised, and Reinforcement Learning
- Whether or not they can learn incrementally
  - Batch versus Online/Incrementally
- Comparison of existing data with new data, or detect pattern using training data
  - Instance based vs model-based training

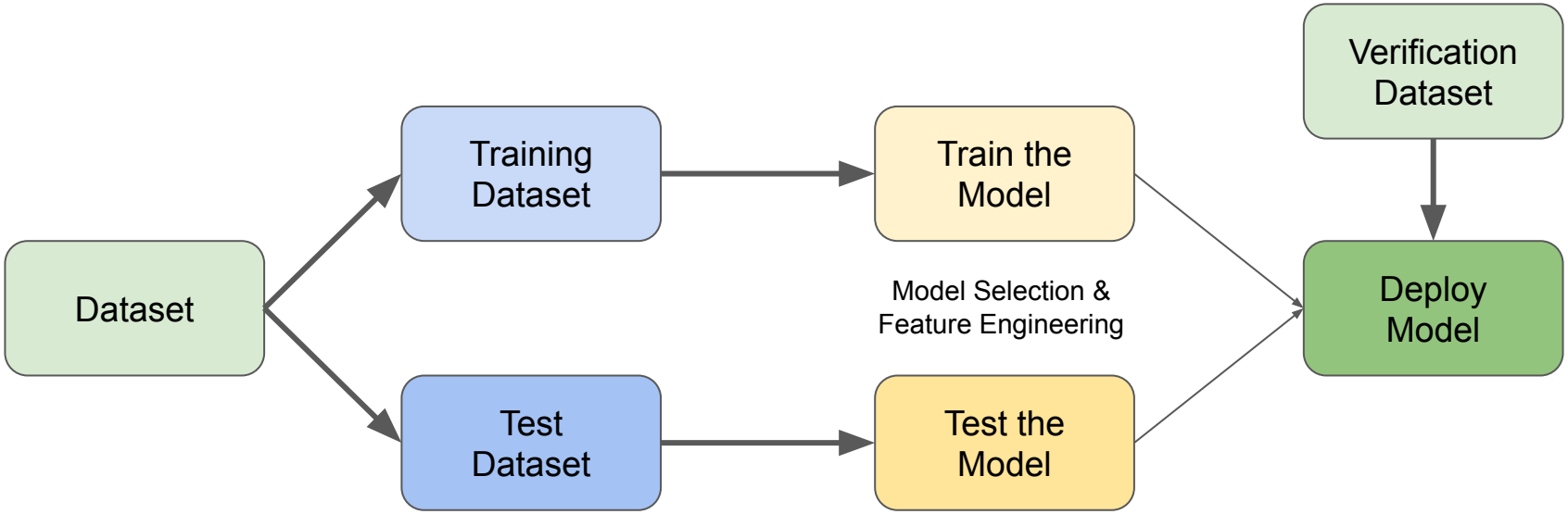
Supervised Learning

In *supervised learning* the training data you feed to the algorithm includes the desired solution called *labels*.

Some examples of supervised learning

- *Classification*: Here the *label*/*target* is one of given set of values. Spam filtering is a good example of this.
- *Regression*: When *target* is a numeric value, and it is continuous in nature (such as car price), then given a set of *features* (mileage, brand, etc) called predictors to predict the target.

WorkFlow - Supervised



ML Academy

SOLIVAR LABS

Supervised Learning Algorithms

- k-Nearest-Neighbors
- Linear Regression
- Logistic Regression
- Support Vector Machines (SVMs)
- Decision Trees and Random Forests
- Neural Networks

Identification of Fruits

ML Academy

SOLIVAR LABS

Five Fruits

ML Academy

SOLIVAR LABS

Types of Apples

ML Academy

SOLIVAR LABS

Supervised Learning

ID	X1	X2	X3	X4	X5	X6	X&	X8	X9	X10	X11	X12	Target

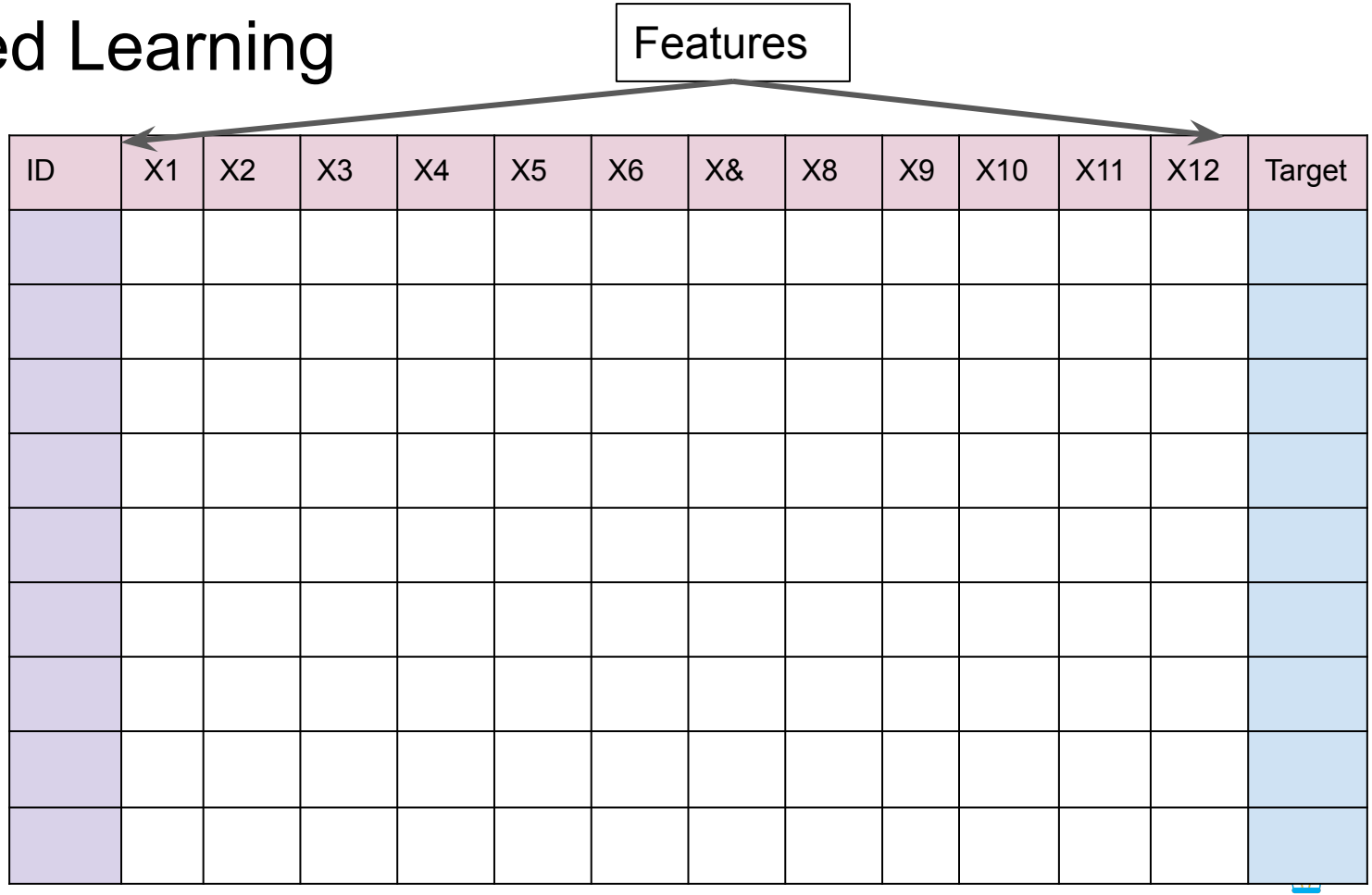
ML Academy

SOLIVAR LABS

Supervised Learning

ID	Size	Color	Shape	Texture	Density	Length	Breadth	Target

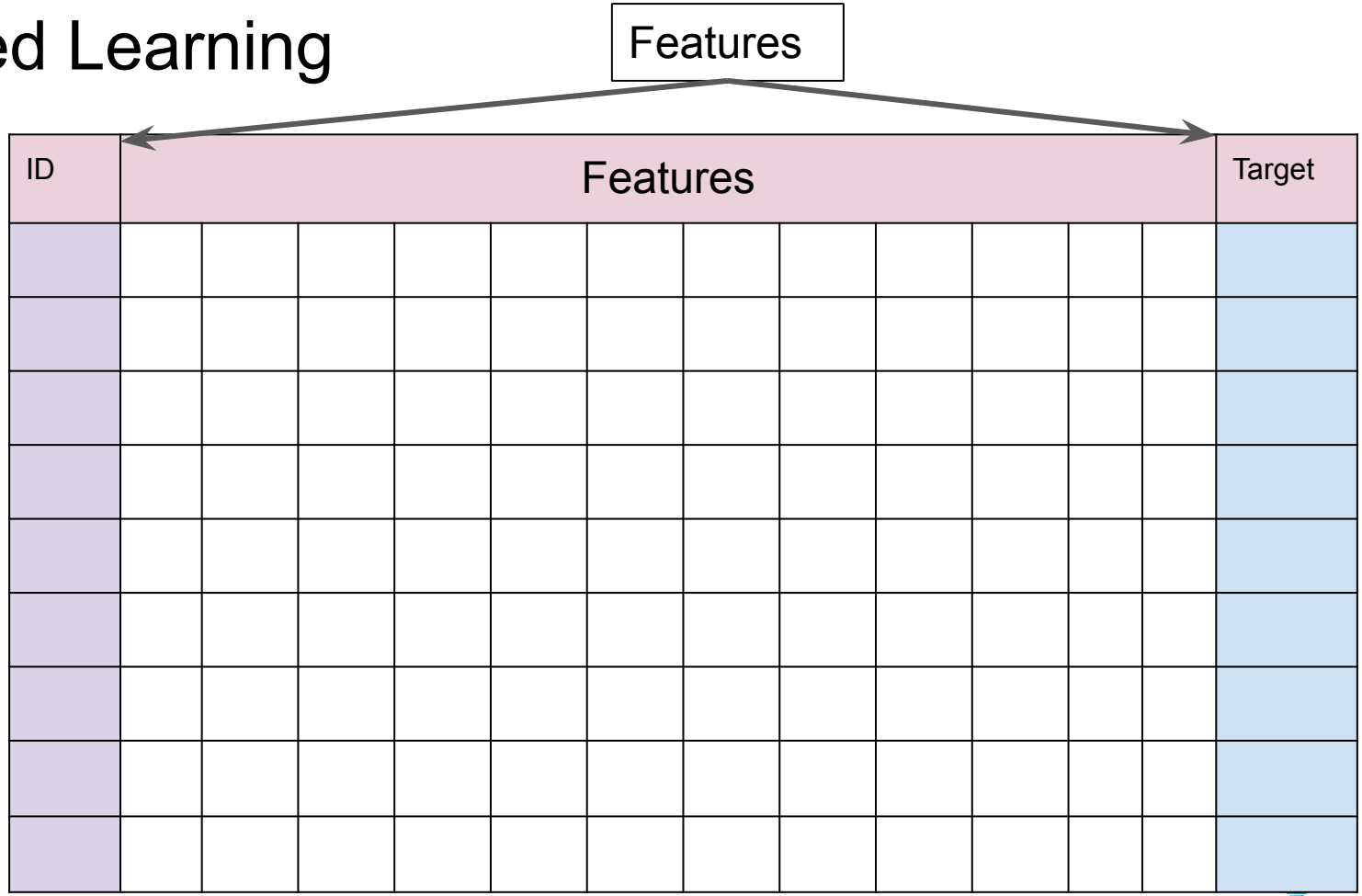
Supervised Learning



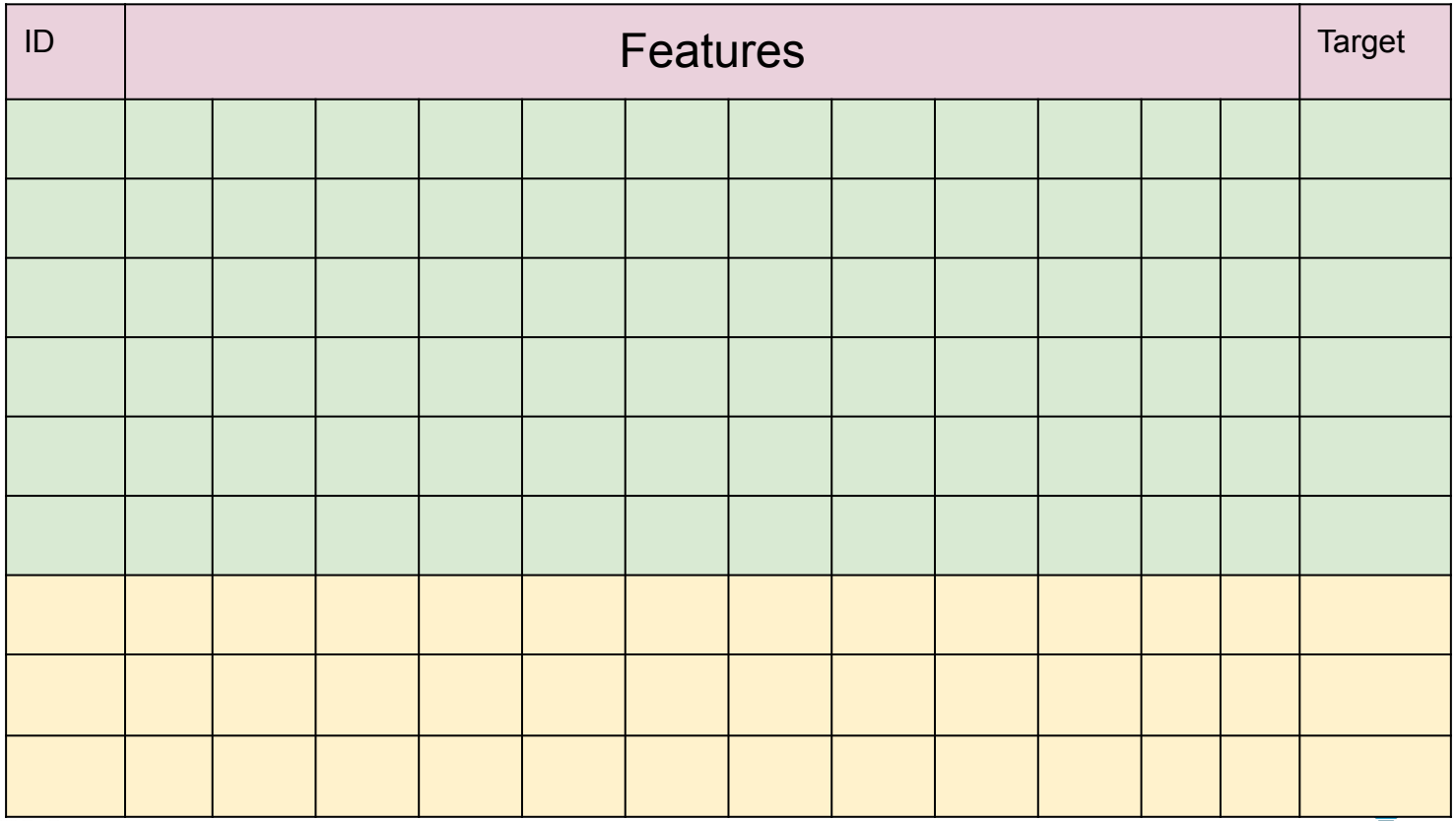
ML Academy

SOLIVAR LABS

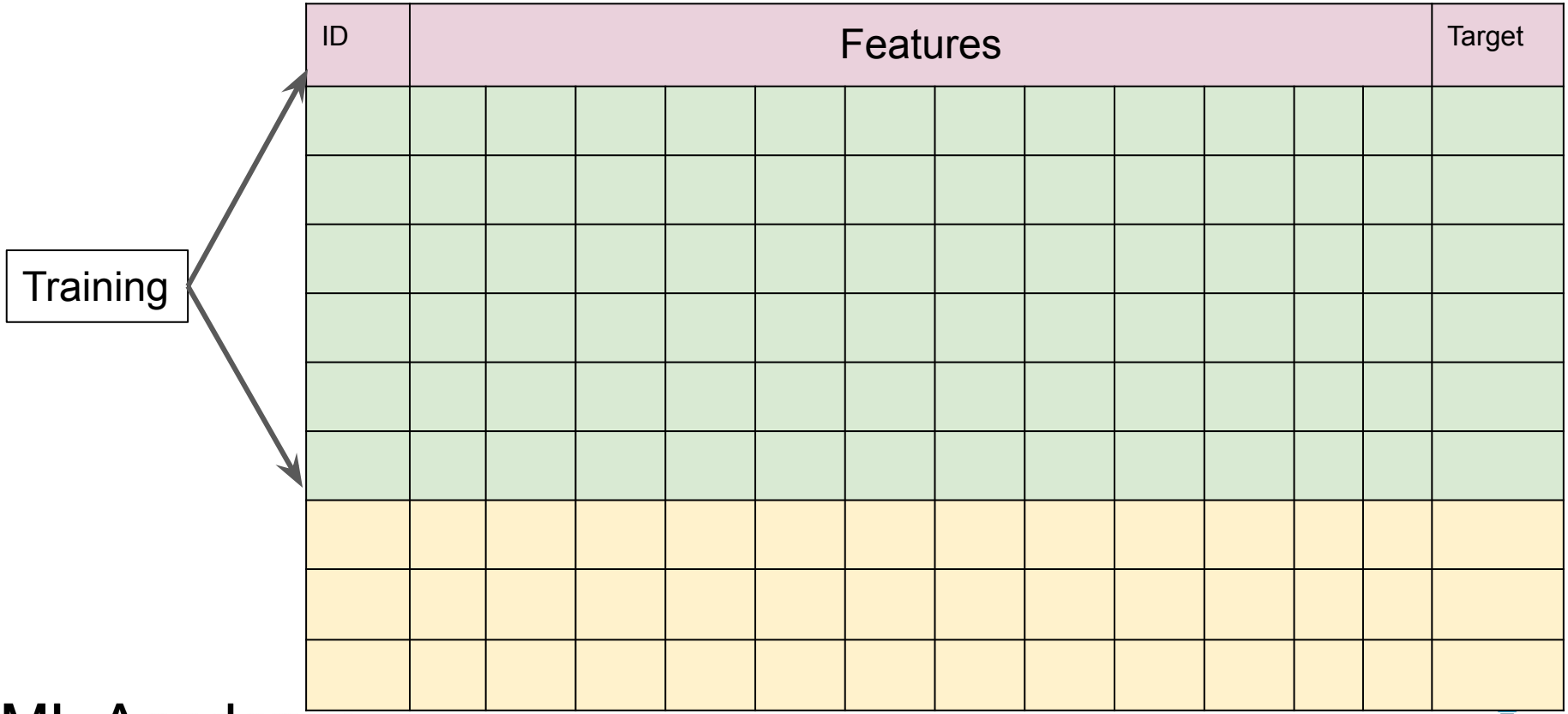
Supervised Learning



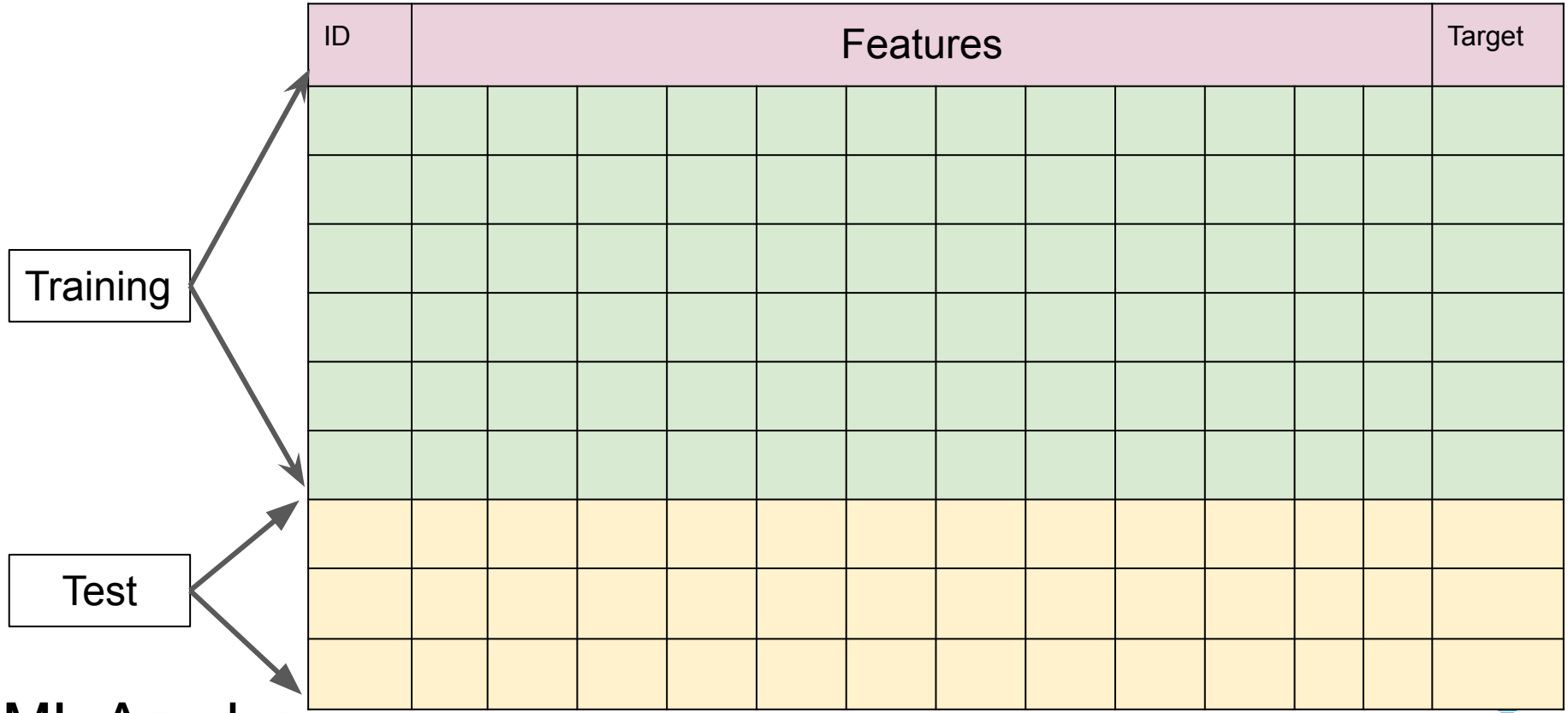
Supervised Learning



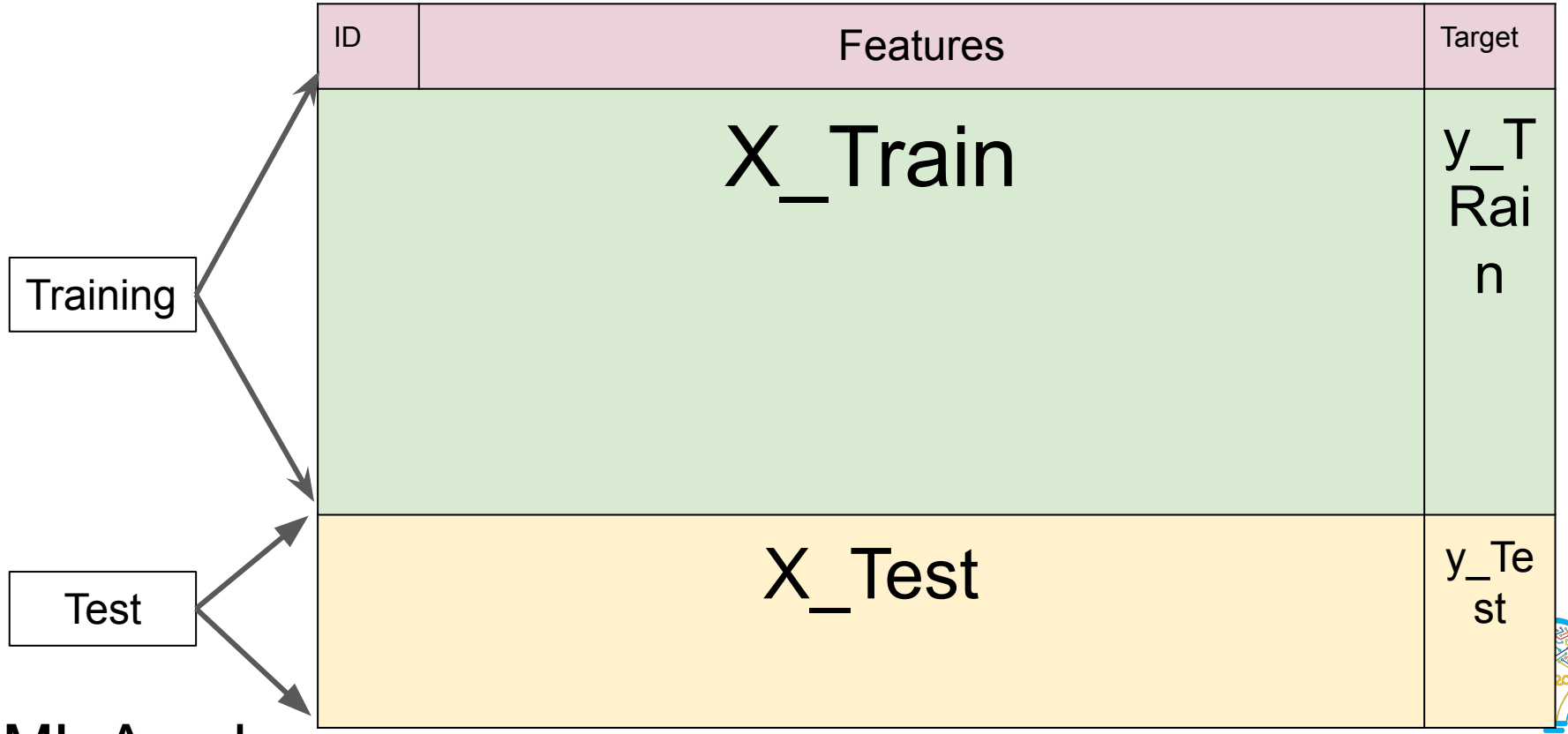
Supervised Learning



Supervised Learning



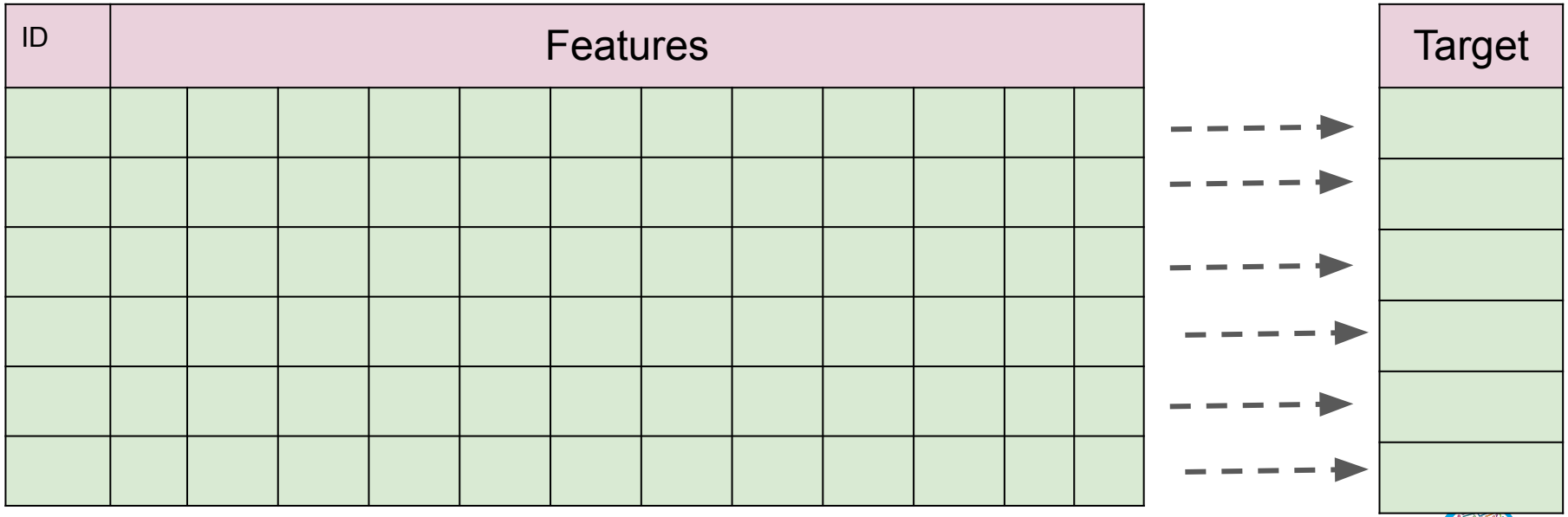
Supervised Learning



ML Academy

SOLIVAR LABS

Some Basic Math



ML Academy

SOLIVAR LABS

Some Basic Math

Target = Function (Features)

Some Basic Math

Target = Fn (X1, X2, X3 - -- - X12)

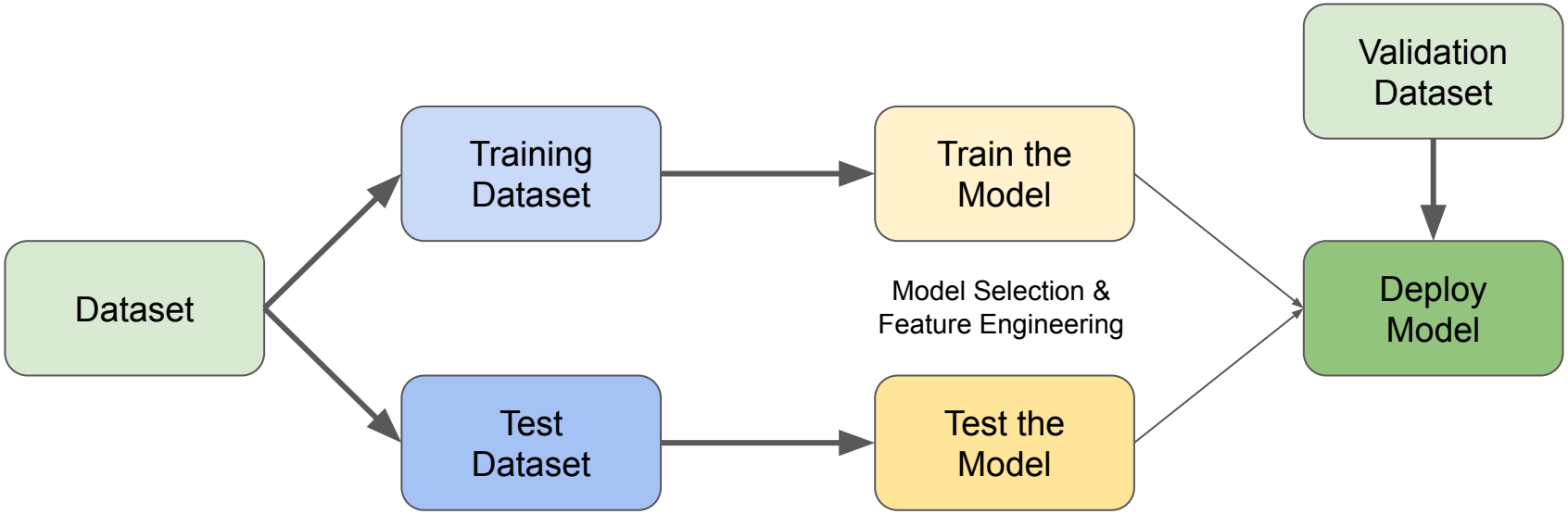
Example of a Linear Function

$$\text{Target} = C0 + C1*X1 + C2*X2 + C3*X3 \; + - -- - + \; C12*X12$$

Machine Learning

- Apply Training set to estimates (C0, C1, C2 …. C12)

WorkFlow - Supervised



ML Academy

SOLIVAR LABS

Supervised Learning

ID	Features											Target	
												Predicted	Actual
												Predicted	Actual
												Predicted	Actual

		Predicted	
		NO	YES
Actual	NO	TN	FP
	YES	FN	TP

ML Academy

SOLIVAR LABS

Confusion Matrix

		Predicted	
		NO	YES
Actual	NO	TN	FP
	YES	FN	TP

Accuracy = (TN + TP)/ (TN + TP + FP + FN)
Precision Score = TP / (TP + FP)

Recall Score = tp / (tp + fn)

F1 = 2 * (precision * recall) / (precision + recall)

ML Academy

SOLIVAR LABS

UnSupervised Learning

ML Academy

SOLIVAR LABS

Unsupervised Learning

The training data is unlabeled. The system tries to learn without a teacher. Example is “Blog visitors categorised by some features”. Some algorithms are:

- Clustering
  - k-Means
  - Hierarchical Cluster Analysis (HCA)
  - Expectation Maximization
- Visualization and Dimensionality Reduction
  - Principal Component Analysis
  - Kernel PCA
  - Locally-Linear Embedding (LLE)
  - t-distributed Stochastic Neighbor Embedding (t-SNE)
- Association Rule Learning
  - Apriori
  - Eclat

Reinforced Learning

- RL is a complete different beast
- The learning system, called in an *agent*, can observe an environment, select and perform actions, and get *rewards* in return. It must then learn by itself what is the best strategy, called a *policy*, to maximize the reward over time.

Batch versus Online Learning

- In batch learning the system is incapable of learning incrementally. It must be trained using all available data.
  - Suggest some examples
- Online / Incremental Learning. In this system you train the system incrementally be feeding data instances sequentially. Either individually or by small groups called mini-batches.
  - Suggest some examples
  - How fast the system can learn is called the *learning rate*.

Instance Based VS Model-Based Learning

Another way to categorize machine learning systems is by how they *generalize*

- **Instance-based Learning**: Learns the examples by heart and then generalizes to new cases using a similarity measure.
- **Model-based Learning**: From a set of examples is to build a model of these examples, then use that model to make predictions.

Main Challenges of Machine Learning

- Insufficient Quantity of training data
- Non Representative of training data
- Poor-Quality data
- Irrelevant Features
  - Critical part of the success Machine Learning project is coming up with a good set of features to train on. This process is called *Feature Engineering*.
    - ***Feature Selection***: selecting the most useful features to train on among existing features.
    - **Feature Extraction**: Combining existing features to produce a more useful one.
- Overfitting the training data
  - Model performs well on training data but does not generalize well.
  - Constraining the model to make it simpler and reduce the risk of overfitting is called *regularization*.

Hands-on Machine Learning Workshops

Goto http://www.mlworkshops.com to find our more about hands-on machine learning workshops and register for upcoming courses under the guidance of Venkatesh Tadinada (CEO, Solivar Labs).

Venkatesh Tadinada
CEO, Solivar Labs

ML Academy

SOLIVAR LABS